INTERCOMPARISON OF THE ECMWF, MSC, AND NCEP GLOBAL ENSEMBLE FORECAST SYSTEMS

Zoltan Toth[1], Roberto Buizza[2], P. L. Houtekamer and Gerald Pelleri[3], Mozheng Wei[4], Yuejian Zhu[5]
[1]NOAA/NCEP/EMC, [2]European Centre for Medium-Range Weather Forecasts, [3]Meteorological Service of Canada, [4]UCAR Visiting Scientist at NOAA/NCEP/EMC, [5]NOAA/NCEP/EMC

This presentation will summarize the methodologies used at the European Centre for Medium-Range Weather Forecasts (ECMWF, Buizza et al. 2003), the Meteorological Service of Canada (MSC, Houtekamer et al. 1996), and the National Centers for Environmental Prediction (NCEP, Toth and Kalnay 1997) to simulate the effects of initial and model uncertainties in ensemble forecasting. In a chaotic system like the atmosphere, probabilistic information is recognized as the optimum format for weather forecasts both from a scientific and a user perspective. Ensemble forecasts are well suited to support the provision of such probabilistic information. In fact, ensembles not only improve forecast accuracy in a traditional sense (by reducing errors in the estimate of the first moment of the forecast probability distribution), but they also offer a practical way of measuring case dependent variations in forecast uncertainty (by providing an estimate of the higher moments of the forecast probability density function). Ensemble forecasting has gained substantial ground in numerical weather prediction over the past decade. Today, many numerical weather prediction centers use ensemble methods in their modelling suite (*WMO 2003*).

National Centers for Environmental Prediction, Washington D. C., US (www.emc.ncep.noaa.gov)
European Centre for Medium-Range Weather Forecasts, Reading UK (www.ecmwf.int)
Meteorological Service of Canada, Dorval, Quebec, Canada (www.msc-smc.ec.gc.ca)
UCAR Visiting Scientist, NCEP/EMC, Washington D.C., US

In this presentation, ensemble techniques (such as the singular vector, multiple analysis cycle, and breeding methods to generate initial perturbations) and the stochastic perturbation and multiple model version techniques for representing model related uncertainty are reviewed and compared. To assess the merit of different existing approaches, operational ensemble forecasts generated at three numerical weather prediction centers were comparatively verified over a 3-month period, May-June-July 2002. Since NCEP generates only 10 perturbed forecasts from each initial time, the quantitative analysis has been limited to 10-member ensembles for ease of comparison and interpretation (the reader should be aware that this induces an under-estimation of the actual skill of the ensemble systems, especially for systems with a large membership, *Buizza & Palmer* 1998).

The main conclusions of the study are that:

The performance of ensemble prediction systems strongly depends on the quality of the data assimilation system used to create the unperturbed (best) initial condition and the numerical model used to generate the forecasts;

A successful ensemble prediction system should simulate the effect of both initial and model related uncertainties on forecast errors; and

For all three global systems, the spread of ensemble forecasts are insufficient to systematically capture reality, suggesting that none of them is able to simulate all sources of forecast uncertainty.

As for some of the detailed results of this study, most verification measures indicate that the ECMWF ensemble forecast system has the best overall performance, with the NCEP system being competitive during the first, and the MSC system during the last few days of the 10-day forecast period. These verification methods, however, measure the overall accuracy of ensemble forecasts influenced by the quality of the data assimilation, numerical weather prediction modelling, and ensemble generation schemes. The results therefore are not directly indicative of the strengths/weaknesses of the different *ensemble generation schemes*. When the forecasts are evaluated using a new technique (PECA) that measures the correlation between forecast error patterns and ensemble perturbations (instead of the full forecasts, thus eliminating the effect of the quality of the analysis on the scores), the overall performance of the three ensemble systems are found to be rather similar.

From a careful analysis of the results based on a comparison of 10-member ensemble systems for May-June-July 2002, a consensus emerges on the following aspects of the systems:

Overall, the EC-EPS exhibits the most skillful performance when measured by RMS, PAC, BSS and ROC-area measures.

When Perturbation vs. Error Correlation Analysis (PECA, Wei and Toth 2003) is used to measure the correlation between the perturbation and forecast-error patterns, the EC-EPS does not show any superior performance. At short lead times, the error patterns are best described by the NCEP-EPS (except on the largest scales), while the MSC-EPS shows the best performance over the largest scales.

Results suggest that the superior skill of the EC-EPS may be mostly due to its superior model and data-assimilation systems, and should not be considered a proof of the superior performance of SV-based initial perturbations. In other words, at MSC and NCEP ensemble performance is negatively affected in the short range by the relatively low quality of the data-assimilation systems, and in the long-range by the relatively low model resolution.

As for statistical reliability, the superior outlier statistics of the MSC-EPS may be due to the use of multiple model versions. This technique may capture large-scale model-related errors for longer lead times.

The spread in the (single-model) EC-EPS grows faster than that in the other two systems due to the stochastic simulation of random model errors, and perhaps to the use of a more active model. There are indications that the scheme for the stochastic simulation of random model error implemented in the ECMWF-EPS improves the forecast statistical reliability.

During the past decade different ensemble generation techniques received significant attention and underwent substantial refinements. Yet a number of open questions still remain. On-going ensemble related research in the coming years is expected to provide a better understanding of the scientific issues still remaining. The comparison of the performance of the ECMWF, MSC, and NCEP ensemble forecast systems reported in this paper can be considered as a necessary first step toward answering some of the open questions. The relative strengths and weaknesses of the three systems identified in this study can offer guidance for the future development of ensemble forecasting techniques. Continuing future collaboration, where initial ensemble perturbations from the three different systems are introduced to the analysis/forecast system of a selected center in a controlled experiment could potentially provide further useful information, contributing to improved forecast operations.

Buizza, R., & T.N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.*, **126**, 2503-2518.

Buizza, R., D. S. Richardson, & T. N. Palmer, 2003: Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensembles. *Q. J. R. Meteorol. Soc.*, **129**, 1269-1288.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie & H. L. Mitchell, 1996. A system simulation approach to ensemble prediction. *Mon. Wea. Rev.,* **124**, 1225-1242.

Toth, Z. & E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297-3319.

Wei, M., & Z. Toth, 2003: A new measure of ensemble performance: perturbation versus error correlation analysis (PECA). *Mon. Wea. Rev.*, **131**, 1549-1565.